Commun. Math. Phys. 271, 431–454 (2007)
Digital Object Identifier (DOI) 10.1007/s00220-007-0201-y

Communications in
Mathematical
Physics

Spin-Flip Dynamics of the Curie-Weiss Model: Loss of Gibbsianness with Possibly Broken Symmetry

Christof Külske[1], **Arnaud Le Ny**[2]

[1] Department of Mathematics and Computing Sciences, University of Groningen, Blauwborgje 3, 9747 AC Groningen, The Netherlands. E-mail: kuelske@math.rug.nl
[2] Université de Paris-Sud, Laboratoire de Mathématiques, Bâtiment 425, 91405 Orsay Cedex, France. E-mail: arnaud.leny@math.u-psud.fr

Received: 14 October 2005 / Accepted: 14 September 2006
Published online: 8 February 2007 – © Springer-Verlag 2007

Abstract: We study the conditional probabilities of the Curie-Weiss Ising model in vanishing external field under a symmetric independent stochastic spin-flip dynamics and discuss their set of points of discontinuity (bad points). We exhibit a complete analysis of the transition between Gibbsian and non-Gibbsian behavior as a function of time, extending the results for the corresponding lattice model, where only partial answers can be obtained. For initial temperature $\beta^{-1} \geq 1$, we prove that the time-evolved measure is always Gibbsian. For $\frac{2}{3} \leq \beta^{-1} < 1$, the time-evolved measure loses its Gibbsian character at a sharp transition time. For $\beta^{-1} < \frac{2}{3}$, we observe the new phenomenon of symmetry-breaking in the set of points of discontinuity: Bad points corresponding to non-zero spin-average appear at a sharp transition time and give rise to biased non-Gibbsianness of the time-evolved measure. These bad points become neutral at a later transition time, while the measure stays non-Gibbs. In our proof we give a detailed description of the phase-diagram of a Curie-Weiss random field Ising model with possibly non-symmetric random field distribution based on bifurcation analysis.

1. Introduction

Recent years have seen a variety of situations where non-Gibbsian lattice-spin measures appear from proper Gibbs measures that are subjected to natural transformations [6, 11]. One particularly interesting dynamical phenomenon is the loss (and possible recovery) of the Gibbs property of an initial Gibbs measure that evolves according to a stochastic spin-flip dynamics. This was analyzed by van Enter *et al.* in [5] and related results for continuous spins that evolve diffusively starting from an initial Gibbs measure were recently obtained in [17]. Suppose a lattice spin system is in an equilibrium situation at initial temperature β^{-1}, described by a proper Gibbs measure in infinite volume. Suppose now the system undergoes a fast heating procedure, described by a high-temperature Glauber dynamics. A question arises: Is there a well-defined Hamiltonian for every time? Or, in the more catchy terms of [5]: Is there always a well-defined temperature? It turns out

that there can be transition-phenomena of an 'in and out of Gibbsianness' as a function of time. These phenomena are already present when we consider an infinite-temperature time evolution, i.e. a spin-flip that is independent from site to site. Let us highlight one particular phenomenon from [5]. Pick the plus-state of a low temperature Ising model in zero external magnetic field as initial measure μ_β at time $t = 0$ and perform an unbiased spin-flip dynamics, independently over the sites, with rate 1. The following result is proved in [5] (Theorem 5.2) for the resulting probability measure $\mu_{\beta,t}$.

Theorem 1.1 ([5]). *Assume that the initial inverse temperature β is above the critical inverse temperature of the nearest neighbor Ising model. Then there exist $t_0(\beta) \leq t_1(\beta)$ such that*

1. *$\mu_{\beta,t}$ is a Gibbs measure for all $0 \leq t < t_0(\beta)$.*
2. *$\mu_{\beta,t}$ is* **not** *a Gibbs measure for all $t > t_1(\beta)$.*

There remain open questions in this picture: Is the transition between Gibbsianness and non-Gibbsianness sharp, i.e. $t_0(\beta) = t_1(\beta)$, as conjectured in [5]? Can one understand the trajectory of the interactions of the measure $\mu_{\beta,t}$ as time varies?

These and related questions seem to be difficult to answer for lattice systems, even for the infinite-temperature dynamics. The purpose of this paper is to provide detailed answers for the corresponding mean-field model. In the course of our analysis we also find an interesting new mechanism of non-Gibbsianness that was not observed on the lattice. Recall that a measure μ on the lattice is Gibbs iff (a version of) its conditional probabilities $\mu(\sigma_x| \cdot) : \sigma_{x^c} \mapsto \mu(\sigma_x|\sigma_{x^c})$ is continuous w.r.t. the product topology for any site x of the graph. This means that the influence of a perturbation of a conditioning σ_{x^c} outside of a large volume Λ tends to zero, when Λ tends to $\mathbb{Z}^d$.

To investigate this property for the time-evolved measure $\mu_{\beta,t}$, we relate its conditional probabilities to expectations w.r.t. a certain *constrained measure*, obtained by conditioning the measure at time $t = 0$ to be in a given spin configuration at time t. In this constrained measure the spin-configuration in the conditioning appears as an additional 'frozen' external magnetic field configuration. The investigation of such a constrained measure was performed in [5] as well as in [17] for the corresponding setup, and we perform it here for a mean-field model.

The failure of the Gibbs-property for the time-evolved measure occurs if there is a sensitive dependence of the constrained model on the external magnetic field configuration, at certain 'bad configurations'. The chessboard configuration, the simplest example of a 'neutral configuration', serves as such a bad configuration in the proof of the second part of Theorem 1.1 in the large β and large t region. In turn, the absence of such bad configurations implies the Gibbs property. To get a precise understanding of the constrained model is a difficult task for lattice models, where one usually cannot hope for exhaustive results. In the mean-field set-up, one can however relate Gibbsianness to continuity properties of conditional probabilities of an infinite-volume constrained model, that can be computed in terms of the rate-function for a standard quenched disordered model. To understand the structure of its minimizers is then equivalent to understand the phase-diagram of a Curie-Weiss random field Ising model with possibly non-symmetric random-field distribution of the quenched disorder, as we shall see.

In this way, our analysis of the transitions between Gibbsian and non-Gibbsian behavior is reduced to the analytic problem of a bifurcation-analysis for this rate-function. We perform this analysis in the spirit of catastrophe theory and discover an interesting new structure where bad configurations (called points in our mean-field approach, see below) with non-zero spin-average appear in a spin-flip invariant model. We call this phenomenon *biased non-Gibbsianness*, as opposed to the more standard *neutral*

non-Gibbsianness where the bad points correspond to configurations with zero spin-average. It is a symmetry breaking in the set of bad configurations that happens in a regime of low enough temperatures, strictly smaller than the regime of temperatures for which there is a phase transition for the initial measure. It corresponds to the region below the so-called *tricritical temperature* of the Curie-Weiss random field Ising model where phase transition is possible for non-symmetric versions of this model. We can find a parametrization in terms of explicit functions of the boundary curve in the (β, t)-space of the region for which $\mu_{\beta,t}$ is Gibbs. This curve consists of two different branches with different functional forms, depending whether they form the boundary to *biased non-Gibbsianness*, where spin-flip symmetry is broken during the loss of Gibbsianness, or to *neutral non-Gibbsianness*.

Biased non-Gibbsianness has not been detected in other models yet. Can a similar property, i.e. non-Gibbsianness with symmetry breaking of bad configurations, also happen for a lattice model? We expect heuristically that the same phenomena should be present also in a Kac model at long enough (but finite) range, see also [7].

2. Model, Results and Outline of the Proofs

2.1. Synopsis and strategy. The model and the time-evolved measure are defined in Sect. 2.2 and the new notions of Gibbsianness and non-Gibbsianness for mean-field models in Sect. 2.3. For lattice systems a Gibbs measure in infinite-volume is characterized by the fact that its conditional probabilities depend continuously on the conditioning, in the sense of the product topology. Clearly, this notion would yield meaningless answers for mean field models, since the finite-volume Gibbs measures of mean field models converge to convex combinations of product measures, by de Finetti's theorem. Non-trivial convex combinations of product measures are known to have conditional probabilities that have every point as a point of discontinuity [8] and thus non-Gibbsianness in this naive approach would always be trivially true for mean-field models.

Observing that mean-field models are first defined at finite volume without boundary conditions, we consider the conditioning at this level first, prescribing the magnetization outside (finite) subsets by *points* α of the interval $[-1, 1]$, before performing an infinite-volume limit. By exchangeability, such a limit of constrained probabilities, if it exists, depends only on this empirical average; thus, good or bad *points* of the interval $[-1, 1]$ arise naturally instead of (spins) configurations on the countable lattice. Similarly, the product topology has to be replaced by the standard Borel topology on $[-1, 1]$. Our approach is motivated by two main observations: Recent progress [9, 11, 16] on generalized Gibbs measures has highlighted the importance of continuity properties for Gibbs measures, and secondly prescribing the values of the magnetization in the conditioning has to do with the macroscopic character of the mean-field interaction. This strategy has moreover already been used with satisfactory results in [10, 15].

Correspondingly, we say that a measure is *Gibbs* if its limiting conditional probabilities are continuous at any point of prescribed outer magnetization, i.e. if all its points are good, *neutral non-Gibbs* if 0 is the only bad point, and *biased non-Gibbs* if there exists a non-zero bad point of prescribed outer magnetization (Definition 2.1). We are ready to state our main results on Gibbsianness during the time-evolution of our model in Sect. 2.4. First, we describe the three possible scenarii depending on the temperature in Theorem 2.2. The high temperature scenario is the simplest one: The time-evolved measure always stays Gibbs. For an intermediate range of temperatures, the phenomenon of neutral non-Gibbsianness appears at a sharp transition time depending on the

temperature only, and stays neutral non-Gibbs forever. The low temperature scenario, below the tricritical temperature of the Curie-Weiss random field Ising model, is more peculiar: After a period of Gibbsianness, biased non-Gibbsianness appears first at a sharp transition time, and becomes neutral at a later sharp transition time.

Transition times as functions of the temperature of the initial measure are given in Theorem 2.3. Biased non-Gibbsianness appears below the tricritical temperature of the Curie-Weiss random field Ising model because of the presence of metastable minima. This model corresponds to the constrained model in which a phase transition is required to get non-Gibbsianness as described in Sect. 3. In this section, we compute the limiting conditional probability in terms of the rate function of the constrained model and transform our Gibbsianness problem to an analysis of the structure of minima of this function. Using elements of catastrophe theory in Sect. 4, we give a detailed phase diagram of this model discovering three main regions of different qualitative behavior of the number and nature of these minima.

Region 1 of the phase diagram (with temperature and dynamical strength of the disorder as parameters) where there is uniqueness of the minimizer for all α, i.e. for all random fields distributions, corresponds to the region of Gibbsianness for the time-evolved measure. Uniqueness of the minimizer indeed implies the absence of phase transition in the quenched model, in the sense that the empirical magnetization always weakly converges to the Dirac measure at this minimizer. The unique phase is thus always selected when one varies around any given value of α, so the conditional probabilities are always continuous and the time-evolved measure is Gibbs. This in particular proves short-times Gibbsianness, known in general for lattice systems [18].

Region 2 is the part of the phase diagram where the rate function of the symmetric case $\alpha = 0$ admits two global minimizers. This is the region of phase transition for the symmetric Curie-Weiss random field Ising model where the empirical magnetization weakly converges to a mixture of Dirac measures at these symmetric minimizers: The two phases coexist and approaching $\alpha = 0$ from above or below yields two different limits corresponding to the two phases. The point $\alpha = 0$ appears thus to be a discontinuity point of the conditional probabilities and Region 2 is the region of neutral non-Gibbsianness.

Region 3 is the region of the phase diagram where there exists a point $\alpha \neq 0$ of the prescribed outer magnetization for which two global minimizers of the rate-function coexist. It strictly includes the region of uniqueness where non-zero metastable minima already appear at $\alpha = 0$. Coexistence of two of the minima is then appearing at some other bigger α_c, as we shall see. In the neighborhood of this α_c it is possible to select different phases, similarly to Region 2 and one recovers a discontinuity of conditional probabilities, but at a non-neutral point of prescribed magnetization. This region is thus the region of biased non-Gibbsianness.

We give in Sect. 4 the precise analysis of the phase diagram, in terms of the temperature and of the strength of the field, of non-symmetric Curie-Weiss random field Ising models. The extended phase diagram and the time evolution picture of Gibbsianness are related in Sect. 5.

2.2. The model. We start at time $t = 0$ with the Curie-Weiss Ising model in zero magnetic field at inverse temperature $\beta > 0$ whose finite-volume Gibbs measures on spin-configurations $\sigma_{[1,N]} = (\sigma_i)_{i=1,\dots,N} \in \Omega_N := \{-1, +1\}^N$ are given by

$$\mu_{\beta,N}(\sigma_{[1,N]}) = \frac{\exp\left(\frac{\beta}{2N}\left(\sum_{i=1}^N \sigma_i\right)^2\right)}{Z_{\beta,N}}, \tag{1}$$

where the normalization factor $Z_{\beta,N}$ is the standard partition function. This model shows a phase transition at the critical inverse temperature $\beta = 1$, distinguishing in the infinite-volume limit a region of uniqueness at high temperature ($\beta \leq 1$) from a region of coexistence of a '+'-like and a '-'-like phase at low temperature ($\beta > 1$). The order parameter of the system is the empirical magnetization

$$m = m_N(\sigma) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i,$$

and we call *point* any possible value of its infinite-volume limit, i.e. a point of the interval $[-1, 1]$.

Then we apply a stochastic spin-flip dynamics at rate 1, independently over the sites. Our object of interest is the resulting image measure $\mu_{\beta,t,N}$ given by

$$\mu_{\beta,t,N}(\eta_{[1,N]}) := \sum_{\sigma_{[1,N]}\in\Omega_N} \mu_{\beta,N}(\sigma_{[1,N]}) \prod_{i=1}^{N} p_t(\sigma_i, \eta_i). \tag{2}$$

Here

$$p_t(\sigma_i, \eta_i) = \frac{e^{\eta_i \sigma_i h_t}}{2\cosh h_t}, \quad \text{with} \quad h_t = \frac{1}{2}\log\frac{1+e^{-2t}}{1-e^{-2t}}, \tag{3}$$

is the transition kernel from a spin-value σ_i at time $t = 0$ to a spin value η_i at $t > 0$. The product form in (2) is a special case of a Glauber dynamics for infinite temperature [5, 18, 19]. It is well known that, for fixed N, the time-evolved measure $\mu_{\beta,t,N}$ tends to a spin-flip invariant product measure on $\{-1, +1\}^N$, with $t \uparrow \infty$. We shall however see that the large-N-behavior of the conditional probabilities of the time-evolved measure can be non-trivial, even when the corresponding measure is close to a product measure.

2.3. The notion of Gibbsianness for mean-field models.

Definition 2.1. *A point $\alpha_0 \in (-1, 1)$ is said to be* **good** *if and only if*

1. *The limit*

$$\gamma_{\beta,t}(\eta_1|\alpha) := \lim_{N\uparrow\infty} \mu_{\beta,t,N}(\eta_1|\eta_{[2,N]}) \qquad \textit{whenever} \lim_{N\uparrow\infty} \frac{1}{N}\sum_{i=2}^{N} \eta_i = \alpha \tag{4}$$

   exists for all α in a neighborhood of α_0.
2. *The function $\alpha \mapsto \gamma_{\beta,t}(\eta_1|\alpha)$ is continuous at $\alpha = \alpha_0$.*

A point is bad when it is not good. We call the mean-field model $\mu_{\beta,t,N}$ **Gibbs** *iff every point α is good,* **neutral non-Gibbs** *iff $\alpha = 0$ is the only bad point and* **biased non-Gibbs** *iff there is a bad point $\alpha \neq 0$.*

For more information and discussion we refer the reader to [10, 15]. Note that the initial mean-field model is Gibbs in this sense because its kernel $\gamma_\beta(\eta_1|\alpha) = \frac{e^{\beta\alpha\eta_1}}{2\cosh(\beta\alpha)}$ is a continuous function at every α.

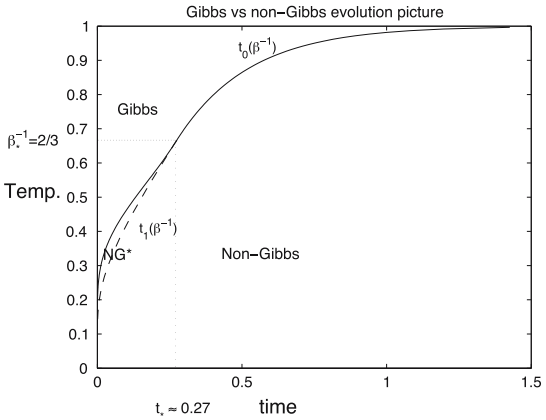

Fig. 1. Gibbs vs non-Gibbs time evolution picture

2.4. Main result. We are now ready to state our main result and describe the Gibbs vs. non-Gibbs character of the time-evolved measure depending on the temperature.

Theorem 2.2.

1. *If $\beta^{-1} \geq 1$, the limiting conditional probabilities are continuous functions of the empirical mean $\alpha \in (-1, 1)$, for all $t \geq 0$ and the time-evolved measure is Gibbs.*
2. *If $\frac{2}{3} \leq \beta^{-1} < 1$, there exists a sharp value $0 < t_0(\beta) < \infty$ such that:*
   - *For all $0 \leq t < t_0(\beta)$, the limiting conditional probabilities are continuous, and the time-evolved measure is Gibbs.*
   - *For all $t \geq t_0(\beta)$, the limiting conditional probabilities are discontinuous at $\alpha = 0$, and continuous at any $\alpha \neq 0$. The time-evolved measure is neutral non-Gibbs.*
3. *If $0 < \beta^{-1} < \frac{2}{3}$, then there exist sharp values $0 < t_0(\beta) < t_1(\beta) < \infty$ such that:*
   - *For all $0 \leq t < t_0(\beta)$, the limiting conditional probabilities are continuous and the time-evolved measure is Gibbs.*
   - *For all $t_0(\beta) \leq t < t_1(\beta)$, there exists $\alpha_c = \alpha_c(\beta, t) \in]0, 1[$, such that the limiting conditional probabilities are discontinuous at the points α_c and $-\alpha_c$, and continuous otherwise. The time-evolved measure is thus biased non-Gibbs and for fixed β^{-1}, the function $t \mapsto \alpha_c(\beta, t)$ is decreasing and we have $\lim_{t \uparrow t_1} \alpha_c(\beta, t) = 0$.*
   - *For $t \geq t_1(\beta)$, the limiting conditional probabilities are discontinuous at $\alpha = 0$, continuous at any $\alpha \neq 0$, and the time-evolved measure is neutral non-Gibbs.*

In this picture (see Fig. 1), we distinguish three regions with different qualitative behavior, one Gibbs region, one neutral non-Gibbs region and one region denoted by "NG*" of biased non-Gibbsianness, due to bad configurations with broken symmetry.

In the next theorem we give the functional form of the two branches of the boundary curves of the Gibbs-region in time-temperature space.

Theorem 2.3.

1. *If $\frac{2}{3} \leq \beta^{-1} < 1$, the threshold time is given by $t_0(\beta^{-1}) = -\frac{1}{4}\log(1 - \beta^{-1})$*

2. *If* $0 < \beta^{-1} < \frac{2}{3}$*, the boundary curve of the Gibbsian region has the parametrization, in terms of a parameter* $M > 0$*,*

$$t_4(M) = -\frac{1}{4}\log\frac{2y^3 + M(1-y^2)^2}{2y + M(1-y^2)},$$
$$\beta_4^{-1}(M) = \frac{y(2+My)(1+y^2)}{2M^2y^3 + 2(y+y^3) + M(1+3y^2+y^4)}, \tag{5}$$

where $y = y(M) = \tanh M$.

We do not have an explicit curve for the line $t_1(\beta)$ but an implicit characterization of the corresponding line of the phase diagram.

In order to understand the symmetry breaking in the set of points of discontinuity in detail, we discuss the *extended phase diagram* of the quenched model, generalizing the known phase diagram from the symmetric case $\alpha = 0$ [1, 21, 3] to non-symmetric distributions. It also has some interest in itself, beyond the study of non-Gibbsianness. Figure 1 then appears as the image of three of the lines of the extended phase diagram, by relating time in the original model to the coupling strength of the quenched fields in the quenched model via $t = \operatorname{atanh}(\exp(-2h_t))$. An explicit expression for the trajectory $t \mapsto \gamma_{\beta,t}$ of the kernel in question in the "Gibbs" region is moreover given in Theorem 3.1. Theorem 2.3 is the result of an analysis of the underlying bifurcation problem.

In Sect. 3 we derive explicit expressions for the limiting conditional probabilities of the time-evolved model in terms of the rate function of a mean-field model with possibly non-symmetric random field distribution. In Sect. 4 we discuss the bifurcation structure of this rate function and the phase diagram of the corresponding model. The proof of Theorems 2.2 and 2.3 follows immediately in Sect. 5 from the phase diagram.

3. Infinite-Volume Limit of Conditional Probabilities

The existence of an infinite-volume limit for the conditional probabilities of the time-evolved model is not granted but must be deduced by an explicit analysis of the model. We show that the limit (4) exists and that a kernel $\gamma_{\beta,t}(\eta_1|\alpha)$ is well-defined when a certain rate-function depending on the triple (β, t, α) has a unique minimizer. In such a case, we have an explicit expression in terms of the minimizer of this rate-function, implying continuity by a smooth dependence of the minimizer on small α-dependent perturbations of the rate-function. On the contrary, when this rate function has more than one minimizer, this dependence is not smooth anymore, leads to different left and right infinite-volume limits and thus gives rise to bad points.

3.1. Continuity result - Explicit expression in terms of a rate-function.

Theorem 3.1. *Denote by* $\tilde{\eta}$ *a random variable taking values* ± 1 *with average* α*, and write* $\mathbb{E}_\alpha$ *for its expectation. Put* $\varepsilon = \varepsilon_t := h_t/\beta$ *and suppose that the function*

$$\Phi_{\beta,\varepsilon,\alpha} : m \mapsto \Phi_{\beta,\varepsilon,\alpha}(m) = \frac{m^2}{2} - \frac{1}{\beta}\mathbb{E}_\alpha\Big(\log\cosh(\beta(m+\varepsilon\tilde{\eta}))\Big) \tag{6}$$

has a unique minimizer $m^*(\beta, \varepsilon, \alpha)$, for the fixed choice of $(\beta, \varepsilon, \alpha)$. Then the conditional probabilities of the time evolved measure (2) have a well-defined infinite-volume limit $\gamma_{\beta,t}$ in the sense of (4) and we have the representation

$$\gamma_{\beta,t}(\eta_1|\alpha) = \frac{e^{\eta_1 g_{\beta,t}(\alpha)}}{2\cosh g_{\beta,t}(\alpha)}, \text{ with } g_{\beta,t}(\alpha) = \frac{1}{2}\log \frac{\cosh\Big(\beta\big(m^*(\beta,\varepsilon_t,\alpha)+\varepsilon_t\big)\Big)}{\cosh\Big(\beta\big(m^*(\beta,\varepsilon_t,\alpha)-\varepsilon_t\big)\Big)}. \quad (7)$$

The (t, α)-dependent effective field $g_{\beta,t}(\alpha)$ captures all information about the interaction at time t of the spin η_1 with the empirical magnetization of the outside world α. The kernel (7) interpolates between the initial distribution and the final product measure. Assuming Theorem 3.1, the further study of discontinuous behavior of the limiting conditional probabilities then boils down to an analytical problem which will be treated in the next section. The result of the theorem itself relies on the following explicit expression of the conditional probabilities in finite-volume in terms of a quenched model.

3.2. Proof: Finite-volume representation in terms of quenched model. For any fixed configuration of 'random fields' $\eta_{[1,N]}$, let us introduce the following 'quenched' measure on the configurations $\sigma_{[1,N]}$:

$$\mu_{\beta,\varepsilon,N}[\eta_{[1,N]}](\sigma_{[1,N]}) := \frac{\exp\Big(\frac{\beta}{2N}\big(\sum_{i=1}^N \sigma_i\big)^2 + \beta\varepsilon\sum_{i=1}^N \eta_i\sigma_i\Big)}{Z_{\beta,\varepsilon,N}[\eta_{[1,N]}]}. \quad (8)$$

This is a Curie-Weiss Ising model with additional random fields η of strength ϵ that will be given by the spin-configuration at time t, constrained to have a fixed empirical distribution α at temperature β^{-1}. The single-site conditional probabilities of the time-evolved model can then be expressed in terms of an expectation w.r.t a quenched model of the form (8) on N-1 spins, depending on $\eta_{[2,N]}$.

Proposition 3.2. *For each finite-volume N we have the following representation of the conditional probabilities of the time-evolved model :*

$$\mu_{\beta,t,N}(\eta_1|\eta_{[2,N]}) = \frac{e^{\eta_1 g_{\beta,t,N}(\eta_{[2,N]})}}{2\cosh g_{\beta,t,N}(\eta_{[2,N]})},$$

where

$$g_{\beta,t,N}(\eta_{[2,N]}) = \frac{1}{2}\log \frac{\mu_{\beta_N,\varepsilon_N,N-1}[\eta_{[2,N]}]\big[\cosh\Big(\beta_N\big(\frac{1}{N-1}\sum_{i=2}^N \sigma_i + \varepsilon_N\big)\Big)\big]}{\mu_{\beta_N,\varepsilon_N,N-1}[\eta_{[2,N]}]\big[\cosh\Big(\beta_N\big(\frac{1}{N-1}\sum_{i=2}^N \sigma_i - \varepsilon_N\big)\Big)\big]} \quad (9)$$

and

$$\beta_N = \beta\cdot\frac{N-1}{N}, \quad \beta_N\varepsilon_N = h_t. \quad (10)$$

Proof. The proof follows from the following rewriting:

$$\frac{\mu_{\beta,t,N}(\eta_1=+|\eta_{[2,N]})}{\mu_{\beta,t,N}(\eta_1=-|\eta_{[2,N]})} = \frac{\sum_{\sigma_{[1,N]}}\exp\Big(\frac{\beta}{2N}\big(\sum_{i=1}^N\sigma_i\big)^2 + \beta\varepsilon\sigma_1 + \beta\varepsilon\sum_{i=2}^N\eta_i\sigma_i\Big)}{\sum_{\sigma_{[1,N]}}\exp\Big(\frac{\beta}{2N}\big(\sum_{i=1}^N\sigma_i\big)^2 - \beta\varepsilon\sigma_1 + \beta\varepsilon\sum_{i=2}^N\eta_i\sigma_i\Big)}.$$

Carrying out the σ_1-sums allows to recognize a quenched model on N-1 spins with frozen random fields $\eta_{[2,N]}$.

Once we have recognized this quenched model, we can reduce its large-N analysis to the determination of the minimizers of a corresponding rate-function. For any fixed configuration $\eta_{[1,N]}$ obeying the constraint $\frac{1}{N}\sum_{i=1}^{N}\eta_i = \alpha$, we use a Gaussian transformation for the empirical σ-sum. The result is standard (see e.g. [2, 14]).

Proposition 3.3. *Let the empirical spin sum $\frac{1}{N}\sum_{i=1}^{N}\sigma_i$ be distributed according to the quenched measure $\mu_{\beta,\varepsilon,N}[\eta_{[1,N]}]$ with random fields obeying the constraint $\frac{1}{N}\sum_{i=1}^{N}\eta_i = \alpha$. Denote by G an auxiliary independent Gaussian $\mathcal{N}(0,1)$ variable, and put*

$$X_N = \frac{1}{N}\sum_{i=1}^{N}\sigma_i + \frac{1}{\sqrt{\beta N}}G.$$

Then the distribution $\mathbb{P}_{\beta,\varepsilon,\alpha;N}$ of the variable X_N has the Lebesgue-density given by

$$\mathbb{P}_{\beta,\varepsilon,\alpha;N}[X_N \in (m, m+dm)] = \frac{e^{-\beta N\Phi_{\beta,\varepsilon,\alpha}(m)}dm}{\int e^{-\beta N\Phi_{\beta,\varepsilon,\alpha}(m)}dm},$$

where the rate function $\Phi_{\beta,\varepsilon,\alpha}$ is given by (6).

The finite-volume distributions are uniquely defined due to permutation invariance. We got an identity for each finite N and the function $\Phi_{\beta,\varepsilon,\alpha}(m)$ is the rate-function for X_N. So we have the convergence of $\frac{1}{N}\sum_{i=1}^{N}\sigma_i$ to the minimizer of Φ when it is unique.

Proof of Theorem 3.1. Put together the explicit representation of Proposition 3.2 of the conditional probabilities with the characterization of Proposition 3.3 of the limit of the quenched model. Then use continuity of the rate-function to get rid of the slight N-dependence (10) and to ignore the influence of the Gaussian variable G.

4. Extended Phase Diagram of Nonsymmetric Random Field Model - Bifurcation Analysis of the Rate-Function

4.1. Description of the extended phase diagram. We discuss now the structure of the minima of the rate function (6) when the parameters β^{-1} (temperature), ε (strength of the random fields), α (asymmetry of the distribution) vary. Recall that β^{-1} is the initial temperature of the system and $\epsilon = \epsilon_t = h_t/\beta$ (3).

The *extended phase diagram* is the decomposition of the parameter-space into regions where the structure of the rate-function does not change. By this we mean that there is no creation of minima (local bifurcations) and also no degeneracy with two global minima. While the potential function (6) depends on three parameters, the parameter α plays a special role to us here. We are interested in values of the parameters $(\beta^{-1}, \varepsilon)$ for which there is a value of $\alpha = \alpha_c$ such that a degeneracy in the depth between two global minima occurs. For the actual computations it is convenient to make a change of variables

$$E = \beta\varepsilon, \quad M = \beta m \quad \text{and write}$$

$$\Psi_{\beta,E,\alpha}(M) := \beta\Phi_{\beta,\varepsilon,\alpha}(m) = \frac{M^2}{2\beta} - \mathbb{E}_\alpha\big[\log\cosh\big(M + E\tilde{\eta}\big)\big]. \tag{11}$$

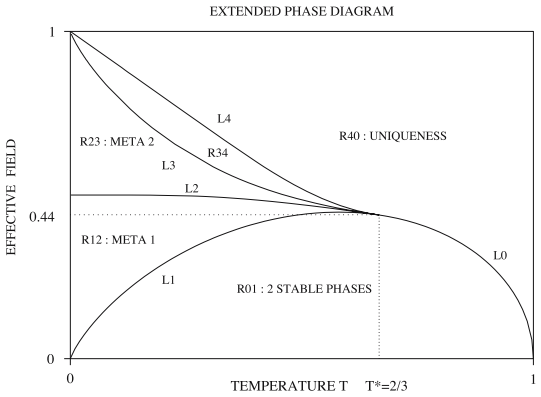

Fig. 2. Extended phase diagram of quenched CW-RFIM

$M \mapsto \Psi_{\beta,E,\alpha=0}(M)$ is an even function that remains unchanged under the joint substitution of (α, M) by $(-\alpha, -M)$, thus we only consider the case $\alpha \in [0, 1]$.

We encounter two mechanisms of creation of new minima, a *fold bifurcation* and a *pitchfork bifurcation* [13, 20]. This analysis has partially been done for the case $\alpha = 0$ [21, 1, 3] and our analysis incorporates these results. There are moreover interesting new phenomena arising when we look at general, possibly non-zero α's. To state this theorem, we write $\bar{\Psi}_{\beta,\epsilon,\alpha} = \Psi_{\beta,\beta\epsilon,\alpha}$ and get the following Fig. 2 describing the different regions of the extended phase diagram :

Region R_{40}: Uniqueness for all α; no phase transition and no degeneracy.
Region R_{01}: Phase transition for $\alpha = 0$; coexistence of two phases, the "+"/"-"-like phases. Uniqueness with possibly metastability for $\alpha \neq 0$.
Region R_{12}: Phase transition with metastability for $\alpha = 0$; coexistence of "+"/"-"-like phases, and a "±"-metastable phase. Uniqueness for $\alpha \neq 0$.
Region R_{23}: Metastability without phase transition; unique neutral "±"-like phase with two "+"/"-"-like metastable phases when $\alpha = 0$. Coexistence for $\alpha_c \neq 0$.
Region R_{34}: Uniqueness with metastability for $\alpha = 0$, degeneracy at some $\alpha \neq 0$ and phase coexistence for $\alpha_c \neq 0$.

Theorem 4.1. *1. If $\beta^{-1} \geq 1$, $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique minimizer M^*.*

2. If $\frac{2}{3} \leq \beta^{-1} < 1$, there exists a sharp value $0 \leq \varepsilon_0(\beta^{-1}) \leq \varepsilon_* = \frac{2}{3} \cdot \operatorname{arctanh}\left(\sqrt{\frac{1}{3}}\right)$:

- *For $\varepsilon < \varepsilon_0$, $\bar{\Psi}_{\beta,\varepsilon,0}$ has two global minimizers $\pm M^*$, $M = 0$ is a local maximizer, but $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique minimizer and no maximizer for all $\alpha \neq 0$.*
- *For $\varepsilon \geq \varepsilon_0$, $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique minimizer M^* for all $\alpha \in [-1, +1]$.*

3. If $0 < \beta^{-1} < \frac{2}{3}$, there exist sharp values

$$0 < \varepsilon_1(\beta^{-1}) < \varepsilon_2(\beta^{-1}) < \varepsilon_3(\beta^{-1}) < \varepsilon_4(\beta^{-1}) \quad \text{s.t.} \tag{12}$$

- *For $\varepsilon \leq \varepsilon_1$, $\bar{\Psi}_{\beta,\varepsilon,0}$ has two minimizers $\pm M^*$ and a local maximizer at $M = 0$; $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique global minimizer and no maximizer for $\alpha \neq 0$.*
- *For $\varepsilon_1 < \varepsilon < \varepsilon_2$, $\bar{\Psi}_{\beta,\varepsilon,0}$ has two minimizers $\pm M^*$, $M = 0$ is a local minimizer and there are two local maximizers $\pm\hat{M}$ with $0 < \hat{M} < M^*$. All minimizers are global for $\varepsilon = \varepsilon_2$ and $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique global minimizer for all $\alpha \neq 0$.*

- *For $\varepsilon_2 < \varepsilon < \varepsilon_3$, $\bar{\Psi}_{\beta,\varepsilon,0}$ has two local minimizers $\pm M^*$, $M = 0$ is the unique global minimizer and there are two local maximizers $\pm\hat{M}$ with $0 < \hat{M} < M^*$.*
- *For $\varepsilon_3 \leq \varepsilon$, $\bar{\Psi}_{\beta,\varepsilon,0}$ has a unique global minimizer $M = 0$ and no local one.*
- *For $\varepsilon_2 < \varepsilon < \varepsilon_4$ there exists $\alpha_c \equiv \alpha_c(\beta^{-1}, \varepsilon) > 0$ s.t $\bar{\Psi}_{\beta,\varepsilon,\alpha_c}$ takes its global minimum at precisely two global minimizers.*
- *For $\varepsilon \geq \varepsilon_4$, $\bar{\Psi}_{\beta,\varepsilon,\alpha}$ has a unique global minimizer, for all α.*

The functions ε_i describe the extended phase diagram (Fig. 2). The regions of the $(\beta^{-1}, \varepsilon)$-plane where the structure of the extrema of the rate-function changes are separated by the lines $(L_i)_{i=0,..,4}$ corresponding to $(\varepsilon_i)_{i=0,..,4}$.

Theorem 4.2 (Characterization of the lines L_i).

1. *There exists a* **tricritical point** $(\beta_*^{-1}, \varepsilon_*)$ *with* $\beta_*^{-1} = \frac{2}{3}$, *where all the lines L_i meet.*
2. *L_0 and L_1 are defined explicitly for $\beta^{-1} \in [0, 1]$ by the same expression for $i = 0, 1$:*

$$\varepsilon_i(\beta^{-1}) = \beta^{-1} \cdot \operatorname{arctanh}\left(\sqrt{1 - \beta^{-1}}\right), \quad \forall \beta^{-1} \in \mathrm{D_i}. \tag{13}$$

3. *L_2 and L_3 can be parameterized by $E > E_* = \beta_*.\varepsilon_1(\beta_*)$ with, for $i = 2, 3$,*

$$\begin{cases} \beta_i^{-1}(E) = G_i(M_i(E), E) \\ \epsilon_i(E) \quad\; = \beta_i^{-1}(E) \cdot E, \end{cases} \tag{14}$$

   with G_i explicitly known, $M_i(E) \neq 0$ implicitly via $F_i(M, E) = 0$, and F_i explicit.
4. *The line L_4 is explicitly parameterized by $M > 0$ in the sense that*

$$\left\{ \begin{pmatrix} \beta^{-1} \\ \varepsilon_4(\beta^{-1}) \end{pmatrix}, 0 < \beta^{-1} < \frac{2}{3} \right\} = \left\{ \beta_4^{-1}(M) \cdot \begin{pmatrix} 1 \\ E_4(M) \end{pmatrix}, 0 < M < \infty \right\} \tag{15}$$

   with

$$E_4(M) = \operatorname{arctanh}\left(\sqrt{\frac{2y^3 + M(1 - y^2)^2}{2y + M(1 - y^2)}}\right), \tag{16}$$

   where $y = y(M) = \tanh(M)$, and $\beta_4^{-1}(M)$ was defined in (5).

The explicit expressions of the F_i's and G_i's entering in the parameterizations of the lines L_i for $i = 2, 3$ are given in the proof (Eqs. (35), (36), (39), (40)). The well-known phase-diagram of the Curie-Weiss random field Ising model with symmetric distribution $\alpha = 0$ was already described in [21]. They obtained L_0 analytically and L_2 numerically. A more complete analysis is required here; we need to know the nature of the extrema as a function of $(\beta^{-1}, \varepsilon, \alpha)$. For the proof of Theorem 2.2 on non-Gibbsianness, only the lines L_0, L_2 and L_4 are relevant, but the study of the lines L_1 and L_3 is valuable to help the understanding of the bifurcation-structure of the rate-function.

4.2. Bifurcation-structure of the rate function. The structure of the extrema of the rate-function is best understood when we fix $E = \beta\varepsilon$ and vary the temperature β^{-1}, so that we are moving along lines of fixed slope E in the phase diagram. There are two fundamentally different cases: $E \leq E_*$ and $E > E_*$,where

$$E_* = \beta_*\varepsilon_* = \operatorname{arctanh}\Big(\frac{\sqrt{3}}{3}\Big) \approx 0.6585$$

corresponds to the slope at the *tricritical point*, whose coordinates are

$$\beta_*^{-1} = \frac{2}{3} \quad \text{and} \quad \varepsilon_* = \frac{2}{3}\operatorname{arctanh}\Big(\frac{\sqrt{3}}{3}\Big) \approx 0.44. \tag{17}$$

This is due to the fact that the fourth derivative of the rate-function for $\alpha = 0$ vanishes at $M = 0$ if and only if $E = E_*$.

(A) Take $E \leq E_*$, assume first $\alpha = 0$ and start from the region above the line L_0 (high-temperature region). Then the rate function has a unique minimizer. When we lower β^{-1} we will eventually cross the line L_0 from the outside. At L_0 we encounter a *pitchfork bifurcation*: The unique minimizer $M = 0$ becomes a local maximizer and two symmetric minimizers are emerging from it. The symmetry is lifted when $\alpha \neq 0$.

(B) Let us now fix $E > E_*$. Put again $\alpha = 0$ first and start from the region below L_0-L_1 where there are two global minima and increase the temperature β^{-1}. Then we first see a pitchfork-bifurcation at the line L_1 that produces a new minimum at $M = 0$ and two local maxima. When we increase β^{-1} this minimum gets deeper to the depth of the non-zero minimum, with equal depth at the line L_2. When we further increase the temperature β^{-1} the zero-minimum becomes the deepest. The non-zero minima will finally vanish at the line L_3 by a *fold bifurcation*, to be discussed below. When we vary α around 0, the degeneracy of the global minima for the non-zero minimizer is lifted in the open region below L_0-L_2. The global minimizer jumps from a positive value to a negative value when we pass $\alpha = 0$ from positive values of α to negative ones.

Suppose we are in the open region between L_2 and L_3. By increasing α from zero, the minimum of the positive minimizer will become deeper w.r.t. the minimum around zero. At the critical value $\alpha_c(\beta^{-1}, \varepsilon)$ the depths of the two minima become equal. This mechanism creates a bad point for the time-evolved measure at a non-zero value of α.

This sort of degeneracy for the specifically chosen $\alpha = \alpha_c(\beta^{-1}, \varepsilon)$ is also possible in the whole open region between L_2 and L_4, even though there is a unique minimizer between L_3 and L_4 for $\alpha = 0$. If we are in particular on the line L_3, increasing α yields a fold bifurcation driven by the parameter α at $\alpha = 0$, for fixed $(\beta^{-1}, \varepsilon)$, creating a new local minimizer for α positive. This also occurs in the region between L_3 and L_4, but at strictly positive values of α. Further increasing the α will then create equal depth of these minima at $\alpha = \alpha_c(\beta^{-1}, \varepsilon)$. The line L_4 is the boundary curve in the parameter space for which this is possible.

4.3. General facts about the rate-function. Note first that $\Psi_{\beta,E,\alpha}(M) \sim \frac{M^2}{2\beta}$ as $|M| \uparrow \infty$ so by continuity and boundedness its infimum is attained at one of its local minima.

Lemma 4.3. *For all (β, E, α), the function $M \mapsto \Psi_{\beta,E,\alpha}(M)$ has at most three local minima. For $\beta^{-1} \geq 1$, it has no maximum and its minimum is the only local extremum.*

Proof. To prove the desired bound of local minima, it suffices to show that the number of zeroes of the second derivative of Ψ (w.r.t. M) is bounded by 4. By an explicit computation, this follows from an elementary discussion of a polynomial in the variable $z = e^{M+E}$. The other statement follows since the curvature is always positive. □

In order to discuss the phase diagram we will need to consider the first four derivatives w.r.t. M. If we denote ∂_M^k the k^{th} derivative with respect to M, one has e.g.

$$\partial_M \Psi_{\beta,E,\alpha}(M) = \frac{M}{\beta} - \mathbb{E}_\alpha\big[\tanh\big(M + E\tilde{\eta}\big)\big],$$

and any minimizer of the rate function must satisfy the so called mean-field equation

$$\frac{M}{\beta} = \mathbb{E}_\alpha\big[\tanh\big(M + E\tilde{\eta}\big)\big]. \tag{18}$$

For any real function $f(M, E)$ we use the notations

$$\bar{f}(M, E) = \frac{1}{2}(f(M+E) + f(M-E)) \quad \text{and} \quad \Delta f(M, E) = \frac{1}{2}(f(M+E) - f(M-E)) \tag{19}$$

for the symmetric average and the discrete derivative around M. For example we write

$$\partial_M \Psi_{\beta,E,\alpha}(M) = \frac{M}{\beta} - \overline{\tanh}\big(M, E\big) - \alpha \Delta \tanh\big(M, E\big).$$

4.4 Elements of catastrophe theory

4.4.1 Catastrophe manifold. To understand the structure of the minima of the *potential function* $M \mapsto \Psi_{\beta,E,\alpha}(M)$ when the three parameters (β, E, α) are varied is an elementary but not completely trivial analytical problem. We find it useful to introduce to this end some systematic basic notions from catastrophe theory going back to Thom and Zeeman (see [20]), to the level that we need them. Catastrophe theory allows for the classification of the possible singularities of local bifurcations of potential functions, as provided by Thom's theorem. It makes rigorous that it suffices to look at polynomial approximations for the potential around critical points to understand locally the nature of bifurcations, i.e. the changes of the structure of the extrema of the potential. In our present case we recognize a *butterfly-singularity* which appears for a four-dimensional family, while we view it here from a three-dimensional sub-space. Indeed, our family has only one parameter α that distinguishes M from $-M$, while in the full butterfly-unfolding one has two such "odd" parameters. As a consequence, not all symmetries are lifted in our present problem.

However, more than just locally identifying the nature of the singularity around a critical point, we manage to provide a fairly explicit analysis over the whole manifold. This is possible by carefully exploiting the specific form of our potential. Taking an elementary but useful message from catastrophe theory, we use M as a coordinate and do not try to solve the appearing equations for M directly. This will be useful throughout the analysis, and allows us in particular to find the explicit parametric form of boundary curve between Gibbsian and non-Gibbsian parts of the phase diagram. Here the magnetization (or state-variable) M appears as a parameter of the curve.

The *catastrophe manifold* X is the set of points in the product space of state space $\mathbb{R}$ and parameter (or *control*) space $C = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ satisfying the mean field equation

$$X = \left\{ (M, \beta^{-1}, E, \alpha) \in \mathbb{R} \times C | \partial_M \Psi_{\beta,E,\alpha}(M) = 0 \right\}.$$

It is a smooth sub-manifold of $\mathbb{R}^4$ that can be characterized by the function

$$\alpha_{\mathrm{MF}}(M, \beta^{-1}, E) := \frac{M\beta^{-1} - \overline{\tanh}\big(M, E\big)}{\Delta \tanh\big(M, E\big)}, \tag{20}$$

well-defined since $E \neq 0$ implies that the denominator does not vanish. Hence X can be written as a graph of the smooth map α_{MF} with the coordinates (M, β^{-1}, E).

The points where this manifold folds "under", when we view the direction of the state variable M as the vertical direction, are points of the vertical tangent plane, and these are given by the vanishing of the curvature of the potential ($\partial_M^2 \Psi_{\beta,E,\alpha}(M) = 0$). The projection to the parameter space of this space, is the *bifurcation set*

$$B := \left\{ (\beta^{-1}, E, \alpha) | \exists M \in \mathbb{R}, \partial_M \Psi_{\beta,E,\alpha}(M) = 0, \partial_M^2 \Psi_{\beta,E,\alpha}(M) = 0 \right\}.$$

It is our aim to describe first this bifurcation set. For our analysis of non-Gibbsianness we are of course interested in the points of global bifurcations where two minima acquire equal depths. To understand this, we show how the analysis of the local bifurcations of the full three-parameter problem can be reduced to a bifurcation problem of a potential function of one variable and the remaining one-dimensional control variable E. We manage here to do this globally and not only near a critical point. To do so we are now going to exploit the linear nature of the potential as a function of the parameters (β^{-1}, α).

Proposition 4.4. *Define the functions*

$$\beta_{12}^{-1}(M, E) := \frac{1 + \tanh(M + E)\tanh(M - E)}{1 + M(\tanh(M + E) + \tanh(M - E))} \tag{21}$$

and

$$\alpha_{12}(M, E) := \frac{M\Big(1 - \overline{\tanh^2}\big(M, E\big)\Big) - \overline{\tanh}\big(M, E\big)}{\Delta\tanh\big(M, E\big) + M\,\Delta\tanh^2\big(M, E\big)}. \tag{22}$$

Then the bifurcation set has the parametric representation

$$B = \left\{ \big(\alpha_{12}(M, E), \beta_{12}^{-1}(M, E), E\big), E > 0, M \in \mathbb{R} \right\}. \tag{23}$$

Remark. Equation (23) gives a parametric representation of the bifurcation set B. The geometry of this set, consisting of a two-dimensional surface with self-intersections, is best understood when we look at intersections of B with planes of fixed E. Performing parametric plots for various fixed values of E, with running variable M, we see a qualitative change of the curves in (β^{-1}, α)-space at $E = E^*$. In going from values $E < E_*$ we see a *cusp-like shape* (Fig. 3) that unfolds into a *pentagram-shaped curve* for $E > E_*$ (Fig. 4).

This singularity is well known from the so-called *butterfly unfolding* (compare [20], from p 178). We note that $\alpha_{12}(M, E) = -\alpha_{12}(-M, E)$ and $\beta_{12}^{-1}(M, E) = \beta_{12}^{-1}(-M, E)$, so the curve is symmetric about the axis $\alpha = 0$. The critical value $E = E^*$ is explained by the fact that $\partial_M^4 \Psi_{\beta^{-1},\alpha=0,E}(M = 0)$ changes sign at $E = E^*$.

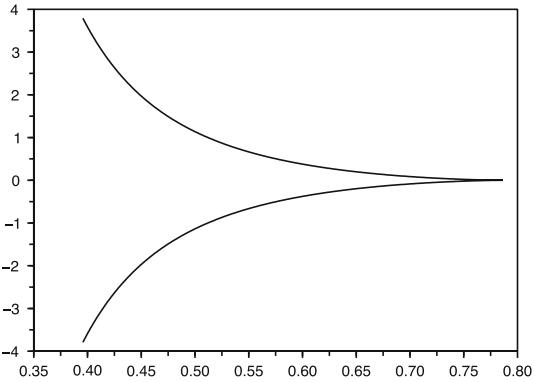

Fig. 3. Cusp: α **vs.** β^{-1} **at fixed** $E = 0.6 < E^*$

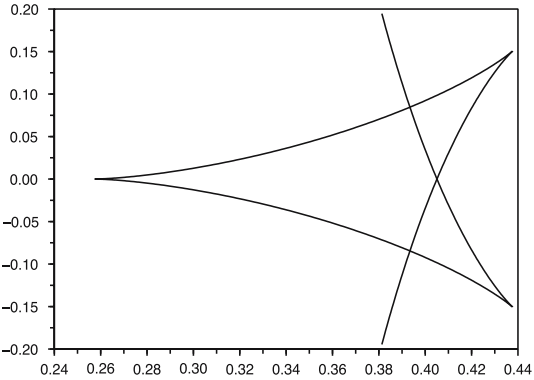

Fig. 4. Pentagram: α **vs.** β^{-1} **at fixed** $E = 1.3 > E^*$

Proof of the Proposition . Consider the equations of the bifurcation set B

$$\begin{cases} \frac{M}{\beta} = \overline{\tanh}\big(M, E\big) + \alpha \Delta \tanh\big(M, E\big) \\ \\ \frac{1}{\beta} = 1 - \overline{\tanh^2}\big(M, E\big) - \alpha \Delta\big(\tanh^2\big)\big(M, E\big). \end{cases}$$

For fixed M and $E > 0$, this is a linear system for β^{-1} and α and the proposition follows from standard arguments showing that it has always a unique solution of the given form $(\beta_{12}^{-1}(M, E), \alpha_{12}(M, E))$.

The point $\big(E_*, \alpha = 0, \beta^{-1} = \frac{2}{3}\big)$ is characterized by the fact that the first four derivatives of the potential at $M = 0$ vanish. Let us discuss the number of minima of the potential function $\Psi_{\beta^{-1},\alpha,E}(M)$ for parameter values outside of B. For $E < E^*$, the potential has a single minimum outside the cusp-shaped region in Fig. 3. Within the cusp, the potential has two minima which are created by a fold bifurcation by crossing a branch of the cusp from the outside. This singularity is the well-known *cusp-catastrophe*, see [20] p. 174.

For $E > E^*$, the situation is more complicated and well-described for instance in Fig. 9.12 in [20]. To the right of the pentagram-shaped curve the potential has a unique

minimum. In the two triangle-shaped regions the potential has two minima. For the upper triangle, one of these minima is created by a fold crossing the upper boundary from above; the other one is created by crossing the right boundary from the left. In the four-cornered region the potential has three minima. Finally, in the region to the left, the potential has two minima. Since the number of minima is a constant for parameter values in the complement of the bifurcation set, we know the number of minima of the potential for any parameter values by extension from the known shape of the potential locally around the critical point, when we know that there are no other types of singularities appearing. That is, we want to prove now that over the whole phase diagram, and not only locally around $(E_*, \alpha = 0, \beta^{-1} = \frac{2}{3})$, this structure persists.

4.4.2. Exact form of L_0, L_1 and L_4 from one-dimensional bifurcation problem. Additionally to fixing E, we fix now the parameter β^{-1}. We want to characterize the points α in the bifurcation set over the (β^{-1}, E)-plane. We will see that there are either 0, 1, 2, 3 or 4 points of such α, and no more, for any fixed (β^{-1}, E), as the plots of the fixed E-slices of B given in Fig. 3 and 4 suggest. We have from the proposition

$$\Big\{\alpha \in \mathbb{R} | (\alpha, \beta^{-1}, E) \in B\Big\} = \Big\{\alpha_{\mathrm{MF}}(M, \beta^{-1}, E), \exists M \in \mathbb{R} : \beta^{-1} = \beta_{12}^{-1}(M, E)\Big\}, \tag{24}$$

where α_{MF} is given by (20). We thus see that by varying the independent parameter β^{-1} we change the number of solutions for M depending on the form of the function $\beta_{12}^{-1}(M, E)$ and the number of its local extrema. The corresponding M gives us a solution $\alpha = \alpha_{12}(M, E)$ on the bifurcation set B. In particular $\beta^{-1} \geq \sup_{M\in\mathbb{R}} \beta_{12}^{-1}(M, E)$ implies that there are no solutions for α, and hence the potential $\Psi_{\beta^{-1},E,\alpha}$ has a unique minimizer, for any α.

To understand the number of solutions for M in (24) we can obtain for different E, we need to proceed now with the bifurcation-analysis of $M \mapsto \beta_{12}^{-1}(M, E)$, viewed as (the inverse of) an auxiliary potential function of the one-dimensional parameter E.

Remark that one can rewrite

$$\beta_{12}^{-1}(M, E) = (1-\tanh^2(E)) \cdot \frac{1+\tanh^2(M)}{1-\tanh^2(M)\tanh^2(E) + 2M\tanh(M)(1-\tanh^2(E))}. \tag{25}$$

Let us denote for $E \leq E_*$,

$$\beta_0^{-1}(E) := \sup_{M\in\mathbb{R}} \beta_{12}^{-1}(M, E), \tag{26}$$

and for $E \geq E_*$,

$$\beta_4^{-1}(E) := \sup_{M\in\mathbb{R}} \beta_{12}^{-1}(M, E). \tag{27}$$

In Fig. 4 the temperature $\beta_4^{-1}(E)$ appears as the projection of the right corners of the pentagram to the β^{-1}-axis. We shall see that these functions indeed define the main critical line L_0-L_4 of the phase diagram whose image to the time-temperature space is in fact the boundary curve of the Gibbsian region given in Fig. 1.

One easily finds $\beta_0^{-1}(E) = 1 - \tanh^2 E$ for $E < E_*$, explaining the functional form (13) of the boundary curve L_0. This functional form for the maximum over M

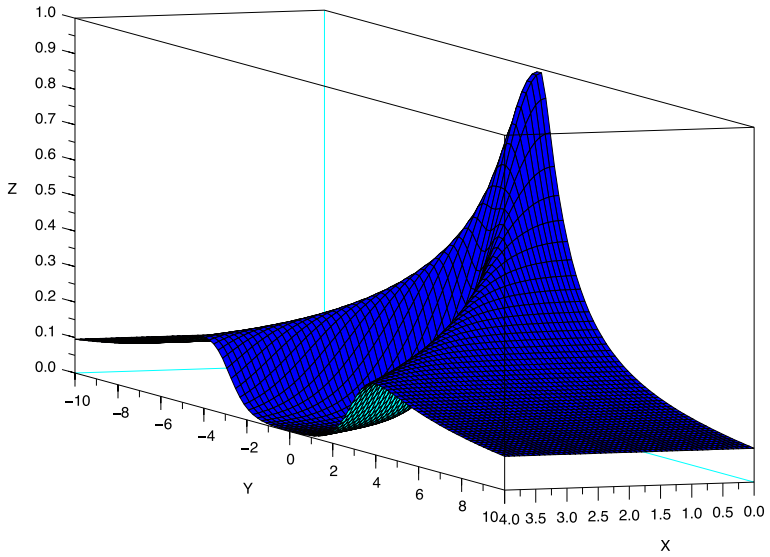

Fig. 5. Pitchfork-bifurcation driven by E at $E = E_*$: potential function $\beta_{12}^{-1}(X = E, Y = M, Z = \beta^{-1})$

is not true anymore for $E > E_*$, as we will see below. Instead we define for $E \geq E_*$ the function $\beta_1^{-1}(E) := \beta_{12}^{-1}(M = 0, E) = 1 - \tanh^2 E$ by this very same functional form as β_0^{-1}. This corresponds to the line L_1 of the phase diagram. This temperature appears as the leftmost corner in Fig. 4. To perform actual computations, we consider the potential $\beta_{12}(M, E)$ with state variable M and control parameter E. Denote by Y the corresponding *catastrophe set*

$$Y = \{(M, E) \in \mathbb{R} \times \mathbb{R}_+ | \partial_M \beta_{12}(M, E) = 0\}.$$

We can find an explicit representation of Y in terms of a smooth curve, writing

$$\beta_{12}(M, E) = \beta_1(E)\Big(U_0(M) - (\tanh^2 E)U_1(M)\Big) \tag{28}$$

with the functions

$$U_0(M) = \frac{1 + 2M \tanh M}{1 + \tanh^2 M} \text{ and } U_1(M) = \frac{\tanh^2 M + 2M \tanh M}{1 + \tanh^2 M}.$$

The following proposition describes the global structure of $\beta_{12}(M, E)$.

Proposition 4.5. *The function $M \mapsto \beta_{12}(M, E)$ undergoes a pitchfork bifurcation where two minima are created by increasing E at the point E_*. More precisely, the catastrophe set has the form $Y = \bar{Y} \cup \Big(\{0\} \times \mathbb{R}_+\Big)$, where $\bar{Y} = \Big\{\big(M, E_4(M)\big), M \in \mathbb{R}\Big\}$ is the graph of the function*

$$E_4(M) = \frac{1}{2} \log \frac{1 + z_4(M)}{1 - z_4(M)}, \tag{29}$$

where $z_4(M) \equiv \sqrt{z_4^2(M)} \geq 0$ *is given by*

$$z_4^2(M) := \frac{U_0'(M)}{U_1'(M)} = \frac{2\tanh^3 M + M(1-\tanh^2 M)^2}{2\tanh M + M(1-\tanh^2 M)} \tag{30}$$

for $M \neq 0$*, and* $z_4^2(M=0) = 0$*. The bifurcation set of (28), the set of E's for which there exists an M s.t.* $\partial_M \beta_{12}(M,E) = \partial_M^2 \beta_{12}(M,E) = 0$*, consists of the single point* $E = E_*$*.*

Remark. $\bar{Y}$ describes a line of global minimizers of $M \mapsto \beta_{12}(M,E)$. Varying the parameter E, we thus have the following picture: If $E \leq E_*$ the function $M \mapsto \beta_{12}(M,E)$ has its unique local minimum at the unique minimizer $M = 0$, and no other local extrema. If $E > E_*$ the minimizer of the function $M \mapsto \beta_{12}(M,E)$ are given by a pair $\pm M_4(E)$, with $M_4(E) > 0$. $M = 0$ is a local maximizer and there are no other local extrema.

Proof. It is easy to show that a pitchfork bifurcation occurs at $E = E_*$. In fact, for this it suffices to show that, at $E = E_*$ and $M = 0$ the derivatives satisfy $\partial_M^2 \beta_{12} = 0$, $\partial_E \partial_M^2 \beta_{12} < 0$ and $\partial_M^4 \beta_{12} > 0$. We skip details here. More interesting is the global analysis. It implies in particular that there are no other bifurcations appearing and the local picture around $E = E_*$ and $M = 0$ extends to the whole space. Note that $\partial_M \beta_{12}(M,E) = 0$ is equivalent to

$$U_0'(M) = (\tanh^2 E) U_1'(M), \tag{31}$$

where the parameter $E > 0$ appears only in terms of multiplication by a function, so that $z^2 = \tanh^2 E$ acts like a re-parameterized linear control variable. It is easy to check that this equation for M always has the solution $M = 0$. Next, we find all the other solutions (after dividing out $M = 0$) by putting

$$\tanh^2 E = z_4^2(M). \tag{32}$$

To show that there are no further bifurcation points other than $E = E_*$, realize that the equations $\partial_M \beta_{12}(M,E) = \partial_M^2 \beta_{12}(M,E) = 0$ imply that $\partial_M z_4^2(M) = 0$. But, given the explicit form (30) of this function of one variable and no parameter, it is an elementary task to show that this implies that $M = 0$. However, from this follows by the second equation $\partial_M^2 \beta_{12}(M=0,E) = 0$ that in fact $E = E_*$ is the only point in the bifurcation set. □

As an important consequence of this we get the desired exact form for L_4.

Theorem 4.6. *The line* L_4 *has the parametric representation of the form*

$$\left\{(\beta_4^{-1}(E), E), E \geq E_*\right\} = \left\{\left(\frac{1 - z_4^2(M)}{U_0(M) - z_4^2(M) U_1(M)}, \frac{1}{2}\log\frac{1+z_4(M)}{1-z_4(M)}\right), M > 0\right\}, \tag{33}$$

where $z_4(M)$ *is given by (30).*

Remark. As a corollary we obtain from here the simple form of L_4 that was given in Theorem 4.2 by a simple computation using the explicit expressions. Note that the parameter M is really β times the magnetization that the system acquires at the melting of the minima at $\alpha_c(\beta, \varepsilon)$.

Proof. Recall that $\beta_4^{-1}(E)$ is the maximum over M of $\beta_{12}^{-1}(M, E)$ for $E \geq E_*$. But from the previous proposition it follows that a *local* maximum is achieved along the curves provided by $\bar{Y}$ in (M, E)-space, for $E \geq E_*$. Since there are no other bifurcations appearing, and the functions go to zero with $|M| \uparrow \infty$ these two symmetric local maxima must automatically be global maxima. So we may parameterize the curve in the (β^{-1}, E)-space in terms of the corresponding M, with e.g. $\beta_4^{-1}(E) = \beta_{12}^{-1}(M, E_4(M))$. From this the representation follows by substitution of the explicit expressions. □

Having understood the form and bifurcation of $\beta_{12}(M, E)$ it is now immediate to go back to the full bifurcation problem and understand the sets of solutions of M for the equation appearing on the r.h.s. of (24):

Corollary 4.7. *Consider the equation $\beta^{-1} = \beta_{12}^{-1}(M, E)$. Fix $E \leq E_*$. Depending whether $\beta_0^{-1}(E)$ is strictly smaller (resp. equal, strictly bigger) than β^{-1}, the number of solutions M is* 0 *(resp.* 1, 2*). Fix now $E > E_*$. Then this number ranges from* 0 *to* 5 *depending on the value of β^{-1} relative to $\beta_1^{-1}(E)$ and $\beta_4^{-1}(E)$.*

The set of α's that lie in the bifurcation set above a point (β^{-1}, E) is given by applying the map α_{12} to the solutions of the above equation. Note that this map is not always injective, indeed the curve $M \mapsto (\alpha_{12}(M, E), \beta_{12}^{-1}(M, E))$ has double-points for $E > E_*$ (we see from the pentagram-shaped Fig. 4 that there are three double-points). Denote by $M_3(E)$, for $E > E_*$, the positive solution of $\alpha_{12}(M, E) = 0$. Then we have $\beta_3^{-1}(E) = \beta_{12}^{-1}(M_3(E), E)$. This is the point of fold-bifurcation at $\alpha = 0$.

4.4.3. Global bifurcations: Degeneracy of minima. Let us finally turn to the discussion of the set of α's such that $M \mapsto \Phi_{\beta,\varepsilon,\alpha}(M)$ has two different global minimizers. We recall that the temperature $\beta_2^{-1}(E)$ is the unique temperature of equal depth of all three minima of the potential at $\alpha = 0$. Therefore it is clear that $\beta_0^{-1}(E) \leq \beta_2^{-1}(E) \leq \beta_3^{-1}(E)$. Exact characterizations of the later two temperatures will be given using different techniques in the next section.

Proposition 4.8. *For any $\beta^{-1} \in (\beta_2^{-1}(E), \beta_4^{-1}(E))$ there is a unique $\alpha_c(\beta^{-1}, E) > 0$ such that there are two different global minima of equal depth. For any other β^{-1} the global minimum is unique for any $\alpha \neq 0$.*

Proof of the Proposition. Fix $E > E_*$. Now, for $\beta = \beta_2(E)$, we have equal depth for the three minima of Φ at $\alpha = 0$. By increasing α, the corresponding minimizers move along smooth curves, until they vanish. The uniqueness of the value of α_c now follows from the following lemma.

Lemma 4.9. *Suppose that, for fixed (β, E), there are two different smooth curves $\alpha \mapsto M_i(\beta, \alpha)$, $i = 1, 2$, α running in the same interval $\alpha \in I$, of local extreme points, where $M_1(\beta, \alpha) < M_2(\beta, \alpha)$. Then the function $\alpha \mapsto \Psi_{\beta,E,\alpha}(M_2(\beta, \alpha)) - \Psi_{\beta,E,\alpha}(M_1(\beta, \alpha))$ is decreasing in α. In particular, there is at most one value of α for which $\Psi_{\beta,E,\alpha}(M_1(\beta, \alpha)) = \Psi_{\beta,E,\alpha}(M_2(\beta, \alpha))$.*

Proof of the Lemma. Looking at the evolution of the values of the rate-function at both curves of extreme points with the parameter α we have

$$\frac{d}{d\alpha}\Psi_{\beta,E,\alpha}(M_i(\beta, \alpha)) = -\frac{1}{2}\log\frac{\cosh(M_i(\beta, \alpha) + E)}{\cosh(M_i(\beta, \alpha) - E)}. \tag{34}$$

It is elementary to see that the function $M \mapsto -\frac{1}{2}\log\frac{\cosh(M+E)}{\cosh(M-E)}$ is decreasing with M (and odd in M), for any fixed $E > 0$. This implies the claim. □

We start now at some point $\beta^{-1} \in (\beta_2^{-1}(E), \beta_3(E)^{-1})$ and $\alpha = 0$: Precisely three local minimizers $\tilde{M}_0 = 0$ and $\tilde{M}_1 = -\tilde{M}_{-1}$ with $\tilde{M}_1 > 0$ exist, with furthermore $\Psi_{\beta^{-1},E,\alpha=0}(\tilde{M}_0) < \Psi_{\beta^{-1},E,\alpha=0}(\tilde{M}_1)$. Indeed, recall that $\beta_2(E)$ is the boundary curve for ferromagnetic order in the phase-diagram of the well-known symmetric Curie-Weiss RFIM. Now, increasing α these minimizers move along curves $\tilde{M}_i(\alpha)$. The lemma implies that $\Psi_{\beta^{-1},E,\alpha}(\tilde{M}_0(\alpha)) - \Psi_{\beta^{-1},E,\alpha}(\tilde{M}_{-1}(\alpha))$ and $\Psi_{\beta^{-1},E,\alpha}(\tilde{M}_1(\alpha)) - \Psi_{\beta^{-1},E,\alpha}(\tilde{M}_0(\alpha))$ decrease, along these curves. Hence $\tilde{M}_{-1}(\alpha)$ plays no role for the discussion of the global minimum because the zero-type minimum is already deeper. It is well-known from the geometry of the butterfly bifurcation and the form of $\Psi_{\beta^{-1},E,\alpha}$ in the pentagram-shaped region that the two left minima are vanishing by fold-bifurcations when increasing α. When the point $M_0(\alpha)$ vanishes by a fold bifurcation, at the corresponding point α^* we have $\Psi_{\beta^{-1},E,\alpha^*}(\tilde{M}_1(\alpha^*)) - \Psi_{\beta^{-1},E,\alpha^*}(\tilde{M}_0(\alpha^*)) < 0$. Hence there is a point $\alpha_c(\beta^{-1}, E)$ when the two right-most minima have equal depth. It is unique because the lemma excludes double crossings.

For the other regions of the phase diagram we use the same arguments. We see in that way that for any other β^{-1} the global minimum is unique for $\alpha \neq 0$. Indeed, consider $\beta^{-1} < \beta_2^{-1}(E)$, $E > E_*$. Then the plus-type minimum is already deeper than the zero-type minimum and becomes even deeper with increasing α, and this excludes the degeneracy of the global minimum.

4.5. The line L_2: degeneracy of non-zero minima and neutral minimum. When $E > E*$, there can be at most two non-negative local minima and it is indeed the case when we cross L_1. When β^{-1} increases, no other local extremum can merge but another type of degeneracy can occur, depending on the depth of the local minima. The region R_{12} between L_1 and L_2 is now characterized by the presence of two symmetric absolute minima at M and $-M$ and a metastable minimum at 0, and all three acquire equal depth at L_2.

We first state that we indeed get a line for L_2 by stating the next lemma which says that degeneracies of the depths of the minima of the 'same type' can only occur for one specific value of β^{-1}, which also provides the sharpness of the transition times of Theorem 2.2.

Lemma 4.10. *Suppose that, for fixed (β, E, α), there are two different smooth curves $\beta^{-1} \mapsto M_i(\beta, E, \alpha)$, $i = 1, 2$, of local extreme points, where $M_1^2(\beta, \alpha) < M_2^2(\beta, \alpha)$. Then the function $\beta^{-1} \mapsto \Psi_{\beta,E,\alpha}(M_2(\beta, E, \alpha)) - \Psi_{\beta,E,\alpha}(M_1(\beta, E, \alpha))$ is decreasing in β^{-1}. In particular, for fixed $E > E_*$, there is at most one value of β^{-1} for which $\Psi_{\beta,E,\alpha}(M_1(\beta, E, \alpha)) = \Psi_{\beta,E,\alpha}(M_2(\beta, E, \alpha))$.*

Proof. The proof is completely analogous to Lemma 4.9, the variation w.r.t. α has only to be replaced by the variation w.r.t β^{-1}. □

So we get the following two equations valid at L_2 for (M, E, β):

$$\begin{cases} \frac{M}{\beta} = \overline{\tanh}(M, E), \\ \\ \frac{M^2}{2\beta} - \overline{\log(\cosh)}(M, E) = -\log(\cosh)(E). \end{cases}$$

Dividing the two equations eliminates the linear variable β^{-1} and implies the following implicit relation of (M, E) characterizing the solutions $F_2(M, E) = 0$ with

$$F_2(M, E) = M\,\overline{\tanh}(M, E) - 2\,\overline{\log\cosh}(M, E) + 2\log\cosh(E). \tag{35}$$

Note that $M = 0$ is always a solution, for any $E \in \mathbb{R}_+$. Looking however for $E \in (E_*, \infty)$, we can numerically find a non-zero solution $M_2 = M_2(E)$. We also emphasize that there can only be one value of $M \neq 0$ for fixed $E > E_*$, since we have proved before that the rate-function can have at most 3 local minima, for any choice of the parameters.

Having solved this (M, E)-relation, the curve L_2 is obtained immediately by solving the mean-field equation for β^{-1}. This gives us the parametric representation

$$E \mapsto \beta_2^{-1}(E) = \frac{\overline{\tanh}(M_2(E), E)}{M_2(E)}. \tag{36}$$

We obtain from this a parametric plot by putting $\varepsilon_2(E) := \beta_2^{-1}(E) \cdot E$.

Let us show that the line has a representation as a function $\varepsilon_2 = \varepsilon_2(\beta^{-1})$, i.e. that $\frac{d}{dE}\beta^{-1}(E) < 0$. To see this we take the derivative of the second equation w.r.t. E and note that the terms involving the M-derivative vanish, due to the mean-field equation. We get

$$\frac{d}{dE}\beta^{-1}(E)\frac{M(E)^2}{2} = \Delta\tanh(M(E), E) - \tanh(E).$$

The proof is now finished by showing that the r.h.s. is strictly less than zero, for all $M > 0$ and $E > 0$. This is seen by writing it in terms of a rational function of the variables $y = e^{2M}$ and $b = e^{2E}$.

4.6. The line L_3: Fold bifurcation. The region between the lines L_2 and L_3 is characterized by a global minimum at $M = 0$ and two symmetric local minima at non-zero values $\pm M^*$, $M^* > 0$ for $\alpha = 0$, non-symmetric for $\alpha \neq 0$. Decreasing β^{-1} along the lines of fixed $E > E_*$ a *fold bifurcation* occurs at the line L_3 : a minimum and a maximum are created from a saddle point by varying the parameter p into one direction. It occurs at a one-dimensional parameter $p = p_0$ whenever

$$\begin{cases} \partial_M \Phi_{p_0}(M_0) = 0, \\ \partial_M^2 \Phi_{p_0}(M_0) = 0, \\ \partial_p \partial_M \Phi_{p_0}(M_0) \neq 0, \\ \partial_M^3 \Phi_{p_0}(M_0) \neq 0. \end{cases} \tag{37}$$

The first equation just says that M_0 is a local minimizer. The second arises because minimizer and maximizer of quadratic type are merging and creating a saddle point.

The line L_3 is thus characterized by the mean field equation at M and furthermore $\partial_M^2 \Psi_{\beta,E,\alpha=0}(M) = 0$. These two equations are reexpressed as

$$\begin{cases} \frac{M}{\beta} = \overline{\tanh}(M, E), \\ \\ \beta^{-1} = 1 - \overline{\tanh^2}(M, E). \end{cases} \tag{38}$$

Again, dividing the two equations eliminates the linear variable β^{-1} and implies an implicit (M, E)-relation of the form $F_3(M, E) = 0$ with

$$F_3(M, E) = M\Big(\overline{1 - \tanh^2}(M, E)\Big) - \overline{\tanh}(M, E). \tag{39}$$

We follow the same logic as for the line L_2. Note that $M = 0$ is always a solution, for any E. For $E \in (E_*, \infty)$ there is a unique solution $M = M_3(E)$ and we can find it numerically. As before, this gives us the parametric representation

$$E \mapsto \beta_3^{-1}(E) = 1 - \overline{\tanh^2}(M_3(E), E) \tag{40}$$

and $\varepsilon_3(E) := \beta_3^{-1}(E).E$. To show that the line has a representation as a function $\varepsilon = \varepsilon_3(\beta^{-1})$, we proceed similarly to L_2.

Let us now vary the parameter α around $\alpha = 0$ at the line L_3. We have

$$\partial_\alpha \partial_M \Psi_{\beta,E,\alpha}(M) = -\Delta \tanh\big(M, E\big) < 0$$

if and only if $M > 0$. But this means that we have a fold bifurcation w.r.t. α, too. The creation of a minimum/maximum-pair is in the direction of increasing α, since the third derivative w.r.t. M of the rate-function has a positive sign. This will allow the extension of the non-Gibbs region to a wider range of temperature for fixed E, namely in the "NG*" region R_{34} between L_3 and L_4.

5. Proof of the Main Results, Conclusions and Perspectives

5.1. Proof of Theorems 2.2 and 2.3. Let us summarize the picture. We are dealing with a quenched system depending on the parameters β^{-1} (initial temperature), ε (time, $\equiv$ strength of the random fields) and α (conditioning of empirical average of spins at time t, $\equiv$ tilting of the random field distribution). Equivalently we are looking at the parameter space spanned by (β^{-1}, E, α). To get the time evolution picture (Fig. 1) from the extended phase diagram (Fig. 2), and thus Theorems 2.2 and 2.3 from Theorems 4.1 and 4.2, we use the smooth relation (3) between time t and $h_t = E_t$. Since $t \longmapsto h_t$ is a decreasing one-to-one map from $[0, \infty)$ to $[0, \infty)$, we get the expressions of the lines of Fig. 1 directly from those of Fig. 2 by inverting (3). The regions of the extended phase diagram with qualitatively different phase structure are mapped into the regions of Gibbsianness, of neutral non-Gibbsianness or of biased non-Gibbsianness.

To prove monotonicity of the critical α_c in NG* as a function of time, we look at the system

$$\begin{cases} \Psi_{\beta,E,\alpha}(M_1) = \Psi_{\beta,E,\alpha}(M_2), \\ \partial_M \Psi_{\beta,E,\alpha}(M_1) = 0, \\ \partial_M \Psi_{\beta,E,\alpha}(M_2) = 0, \end{cases}$$

and show that $\frac{d\alpha}{dE}\Big|_{\beta^{-1}} \geq 0$. Taking the derivative of the first equation w.r.t. E and using the second and the third equation, the claim follows because $M \mapsto \Delta \log\cosh\big(M, E\big)$ is an increasing function and $M \mapsto \mathbb{E}_\alpha \tanh\big(M + E\tilde{\eta}\big)\tilde{\eta}$ is descreasing.

5.2. Conclusions and open questions. In this study of the extension of the Gibbs property from lattice systems to mean-field models, we have recovered the results that correspond to what has been proved on the lattice. Completing the picture, we discovered a new phenomenon, *biased non-Gibbsianness*. This is another indicator of the relevance of the definition of Gibbsianness for mean-field models recently introduced in [10, 15] and used in this paper, but other results and examples would be welcome. It is indeed well known that mean-field results transfer only partially to lattice systems, with sometimes peculiarities (van der Waals theory, equivalence of ensembles, structure of the Gibbs measures, etc.). Not all analogies should be taken too literal: For lattice systems, in $d=2$, the chessboard configuration is bad, while a neutral random configuration is good, but in the mean-field set-up the difference disappears and the conditioning leads to the same constrained measure. Care is thus needed to extend mean-field results to lattices, but they remain a good tool to detect new types of phenomena. These may also be recovered for limiting cases of lattice measures. In particular, it would be interesting to know whether the new phenomenon of biased non-Gibbsianness could occur on lattices. Kac models are good candidates to study this question and a positive answer, even partial, would support the relevance of the notion of Gibbsianness in the mean-field set-up.

Acknowledgements. We thank the Laboratoire de Mathématiques d'Orsay (CK), Eurandom (ALN, CK) and University of Groningen (ALN) for invitations, crucial in the achievement of this work. We thank A. van Enter, R. Fernández, E. Orlandi, F. Redig, M. Vares and V. Zagrebnov for encouraging discussions, H. Broer for advice on catastrophe theory, and we are grateful to L.M. Le Ny for the help in scilab and matlab to draw the pictures.

References

1. Amaro de Matos, J.M.G., Patrick, A.E., Zagrebnov, V.: Random infinite-volume Gibbs states for the Curie-Weiss random field Ising model. J. Statist. Phys. **66**, 139–164 (1992)
2. Bovier, A.: Statistical mechanics of disordered systems. *A mathematical perspective, Cambridge Series in Statistical and Probabilistic Mathematics* **18**,Cambridge: Cambridge University Press, 2006
3. Cassandro, M., Orlandi, E., Picco, P., Vares, M.E.: One-dimensional random field Kac's model: localization of the phases. Elec. J. Probab. **10**, 786–864 (2005)
4. Ellis, R.S.: Entropy, *Large Deviations and Statistical Mechanics*. Newyork: Springer-Verlag 1985
5. van Enter, A.C.D., Fernández, R., den Hollander, F., Redig, F.: Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures. Commun. Math. Phys. **226**, 101–130 (2002)
6. van Enter, A.C.D., Fernández, R., Sokal, A.D.: Regularity properties of position-space renormalization group transformations: Scope and limitations of Gibbsian theory. J. Statist. Phys. **72**, 879–1167 (1993)
7. van Enter, A.C.D., Külske, C.: Two connections between random systems and non-Gibbsian measures. To appear in J. Statist. Phys., DOI 10.1007/s10955-006-9185-9
8. van Enter, A.C.D., Lörinczi, J.: Robustness of the non-Gibbsian property: some examples. J. Phys. A **29**, 2465–2473 (1996)
9. van Enter, A.C.D., Verbitskiy, E.A.: On the variational principle for Generalized Gibbs measures. In: Proceedings of the workshop "*Gibbs vs. non-Gibbs in statistical mechanics and related fields*" (Eurandom, 2003), Markov proc. Relat. Fields **10**(2), 2004
10. Häggström, O., Külske, C.: Gibbs property of the fuzzy Potts model on trees and in mean-field. In: *Proceedings of the workshop "Gibbs vs. non-Gibbs in statistical mechanics and related fields*" (Eurandom, 2003), Markov proc. Relat. Fields **10**(2), 2004
11. Fernández, R.: Gibbsianness and non-Gibbsianness in Lattice random fields. In: *Mathematical Statistical Physics*. Proceedings of the 83rd Les Houches Summer School, July 2005, London: Elsevier, 2006
12. Georgii, H.-O.: *Gibbs Measures and Phase Transitions*, NY, de Gruyter 1988
13. Guckenheimer, J., Holmes, P.: *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Applied Mathematical Sciences, **42**, Springer, NY 1990
14. Külske, C.: Metastates in disordered mean-field models: Random field and Hopfield models, J. Statist. Phys. **88**(5/6), 1257–1293 (1997)
15. Külske, C.: Analogues of non-Gibbsianness in joint measures of disordered mean field models, J. Statist. Phys. **112**, 1101–1130 (2003)

16. Külske, C., Le Ny, A., Redig, F.: Relative entropy and variational properties of generalized Gibbs measures. Ann. Probab. **32**(2), 1691–1726 (2004)
17. Külske, C., Redig, F.: Loss without recovery of Gibbsianness during diffusion of continuous spins. Probab. Theor. Relat. Fields. **135**(3), 428–456 (2006)
18. Le Ny, A., Redig, F.: Short times conservation of Gibbsianness under local stochastic evolutions. J. Statist. Phys. **109**(5/6), 1073–1090 (2002)
19. Liggett, T.M.: *Interacting Particle Systems*. NY Springer-Verlag, 1985
20. Poston, T., Stewart, I.: Catastrophe Theory and its Applications, *Surveys and reference works in mathematics*, London: Pitman (1978)
21. Salinas, S.R., Wreszinski, W.F.: On the mean-field Ising model in a random external field. J. Statist. Phys. **41**(1/2), 299–313 (1985)

Communicated by J.L. Lebowitz